Big Data and Computing Visions

www.bidacv.com

Big. Data. Comp. Vis. Vol. 1, No. 3 (2021) 161–169.

Paper Type: Original Article

Efficient Energy Consumption in Smart Buildings using Personalized NILM-based Recommender System

Fatemeh Taghvaei, Ramin Safa*

Department of Computer Engineering, Ayandegan Institute of Higher Education, Tonekabon, Iran; fatemeh_taghvai@yahoo.com; safa@aihe.ac.ir.

Citation:

Taghvaei, F., & Safa, R. (2021). Efficient energy consumption in smart buildings using personalized NILM-based recommender system. *Big data and computing visions, 1* (3), 161-169.

Received: 08/04/2021 Reviewed: 12/05/2021 Revised: 19/06/2021 Accept: 07/07/2021

Abstract

As the construction sector accounts for the highest energy consumption worldwide, new solutions must be offered in buildings through the adoption of energy-efficient techniques. The main factors involved in energy consumption and residents' behaviors patterns considering environmentally-friendly lifestyle changes must be clearly identified and modeled to provide such solutions. One of the most important topics in smart grids is managing energy consumption in buildings, and one way to optimize energy consumption by analyzing building energy data is to use personalized recommender systems. The Non-Intrusive Load Monitoring (NILM) technique is an important way to cost-effective real-time monitoring the energy consumption and time of use for each appliance. However, the combination of recommender systems and NILM has received less attention. In this paper, a personalized NILM-based recommender system is proposed, which has three main phases: DAE-based NILM, TF-IDF-based text classification, and personalized recommender system. The proposed approach is investigated using the Reference Energy Disaggregation Dataset (REDD). According to the results, the accuracy of the proposed framework is about 60%.

Keywords: Smart buildings, Recommender Systems, NILM, Deep Learning, TF-IDF.

1 | Introduction

Energy consumption in buildings should be optimized based on consumption behaviors due to the reduction of fossil fuels and increasing energy demand. The problem of energy sharing using Non-Intrusive Load Monitoring (NILM) is a new approach to reduce energy consumption in smart homes. Information on energy consumption patterns is essential for energy users and managers to design energy management strategies and customize energy demand. NILM techniques provide this information by distributing electrical charges concentrated at the household level [1]. New approaches to energy efficiency use a combination of recommender systems and NILM to reduce household energy consumption. These approaches use the NILM technique as a load disaggregation to identify the needs and preferences of households and suggest optimal supplies using recommender systems.

Corresponding Author: safa@aihe.ac.ir

10.22105/bdcv.2022.325031.1039

Taghvaei & Safa| Big. Data. Comp. Vis. 1(3) (2021) 161-169

In the field of NILM, machine learning-based techniques have received more attention than the Hidden Markov Model (HMM) and metaheuristic algorithms. The idea proposed in the study on the use of NILM is based on deep learning for a personalized recommender system to optimize energy consumption in smart buildings. In this study, the NILM energy disaggregation technique based on De-noising Auto-Encoders (DAE) is used. Deep learning techniques such as DAE can automatically detect high-level features without the need for user intervention. Since the energy signals aggregated in smart homes often contain some noise, the use of a DAE can detect and eliminate these noises as much as possible and improve the performance of the NILM system. The NILM technique automatically separates smart meter data into different profiles. In the proposed method, the consumption patterns of each household appliance are extracted in a user profile based on the NILM results. At the same time, the personalized recommender system automatically collects advertisements for optimal home appliances and uses information retrieval techniques to represent each one as a home appliance profile. A similarity measurement is then applied to compare the user profiles and home appliance profiles for ranking. Finally, appropriate recommendations are provided to users.

The paper is organized as follows. The second section introduces a variety of energy disaggregation approaches in different categories focusing on machine learning-based NILM. In the third section, a new method is proposed that uses NILM, TF-IDF, and cosine similarity techniques to develop a recommender system. The fourth section presents the results of the simulation and analysis, and the fifth section provides a conclusion from what is discussed in the paper.

2 | Literature Review

The NILM technique was first developed in the 1980s [2]. Energy disaggregation, known in the literature as NILM, is the task of inferring the power demand of the individual appliances given the aggregate power demand recorded by a single smart meter that monitors multiple appliances [3]. Energy consumption in a building is equal to the total electricity consumption of all appliances. Therefore, the amount of electricity consumed by each appliance must be specified. The aggregated energy of N appliances at time t can be the sum of the active energies of all the appliances, which is defined as *Eq (1)*.

$$y(t) = \sum_{i=1}^{N} y_i(t) + e(t) \tag{1}$$

Where $y(t)$ is the amount of energy aggregated at time t, N is the number of appliances, $y_i(t)$ is the share of appliance i in energy consumption at time t, and $e(t)$ is noise or unwanted energy.

Energy data analysis can be divided into three categories: 1) supervised 2) unsupervised, and 3) evolutionary algorithms. Supervised learning methods require labeled samples of each appliance to develop a model. The use of machine learning techniques in NILM was first proposed using neural networks to identify the energy consumption of home appliances [4]. Sparse auto-encoders [5], back-propagation, learning vector quantization [6], and recurrent neural networks [7] were then proposed for energy disaggregation. Despite the difficult training process, deep neural networks perform better than Hidden Markov Models. The use of Long Short-Term Memory (LSTM) architecture and DAEs has recently been proposed to identify defective appliances [8]. Another area of application of neural networks for NILM is the use of the Restricted Boltzmann Machine (RBM) [9, 10] and Convolutional Neural Network (CNN) [10].

Unlike supervised NILM, unsupervised techniques require no pre-training and, consequently, less costly and more reliable for real-time NILM. Unsupervised NILM approaches can be divided into three subgroups: 1) Unsupervised approaches that require unlabeled training to develop appliance models. They are often HMM-based and produce appliance models either manually or automatically during the

training phase [11]; For example, a single hardware setup can be used to measure the energy consumption of each appliance in the household. This hardware setup with the use of certain machine learing algorithms like Factorial Hidden Markov Model (FHMM) and Combinatorial Optimization (CO) disaggregates the combined household energy readings to device specific values. These values then get sent to a cloud database and are presented to the user through a dashboard like visual interface [12]. 2) Unsupervised approaches that use labeled data to develop appliance models and then use models to disaggregate energy in buildings. These approaches require the collection of appliance data. This data is used to develop appliance models that are then used in new buildings. Most deep learning-based NILM techniques fall into this category, and 3) Unsupervised approaches require no training before energy disaggregation. These approaches can disaggregate energy without the need for measured data or prior knowledge.

The genetic algorithm, and evolutionary algorithm, have been also used in some studies on NILM. The genetic algorithm is mainly used to identify the features and patterns of energy profiles of appliances and to optimize existing parameters used in fuzzy systems [13, 14, 15]. The differential evolution algorithm is a simple and robust metaheuristic approach that uses genetic algorithm operators such as crossover, mutation, and selection and is used for NILM [16].

2 | The Proposed Method

Recommender systems are powerful tools for improving energy efficiency that can understand users' preferences and needs and recommend useful energy products/services to each household individually by analyzing their data. This method is proposed to implicitly infer the needs and preferences of appliance users through the analysis of energy consumption patterns and encourage users to purchase appliances that are energy efficient to potentially reduce energy consumption and energy costs and, consequently, provide user satisfaction. The proposed framework for personalized NILM-based recommender system in the smart buildings can be seen in *Fig. 1*. The proposed framework has eight stages: 1) Preprocessing, 2) DAE, 3) Load detection, 4) Creating preference matrix, 5) Generating household profiles, 6) Calculating TF-IDF, 7) Generating item profiles, and 8) Generating recommendation list. We will explain the preprocessing and DAE stages below.

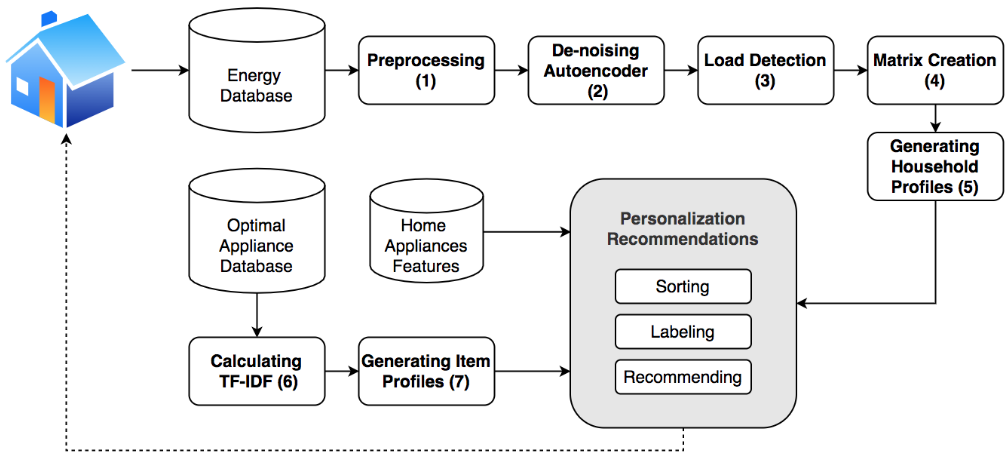


Fig. 1. The proposed framework for personalized NILM-based recommender system

2.1 | Pre-Processing

The pre-processing stage consists of three phases: 1) Normalization, 2) Appliance selection, and 3) Sample selection. In the first phase, all values are normalized between 0 and 1 using min-max normalization. Next, 11 appliances, including microwaves, washing machines, dishwashers, refrigerators, electric stoves, water

heaters, air conditioners, electric heaters, electric ovens, kitchen appliances, and gas stoves, will be selected. In the proposed method, all homes are used for training and testing so that four-day data is used for testing and others are used for training.

2.2 | De-noising auto-encoder

The NILM problem can be formulated as a de-noising problem by expressing the aggregated signal as the total energy consumed by the appliance along with a noise value that makes up the entire remaining value. In particular, *Eq. (2)* could be written as follows.

$$y(t) = y_j(t) + v_j(t) \tag{2}$$

$$v_j(t) = \sum_{\substack{i=1 \\ i \neq j}}^{N} y_i(t) + e(t) \tag{3}$$

Eq. (3) represents the noise for the appliance j. So, the noise $v_j(t)$ must be taken out of the aggregated measurement to obtain $y_j(t)$. In the NILM field, an auto-encoder is fit to the appliance to reconstruct $y_j(t)$ according to the aggregated signal $y(t)$. A DAE is an auto-encoder that tries to reconstruct the noise-free input from the noisy one. *Fig. 2* show the NILM estimated results.

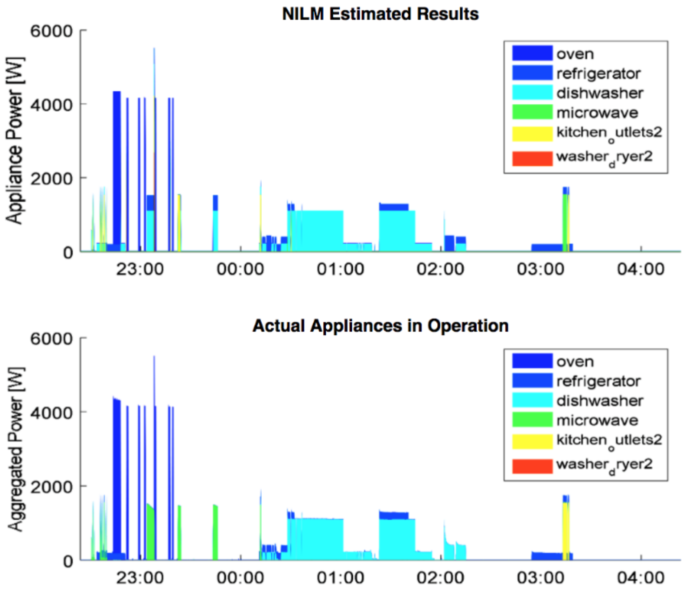


Fig. 2. NILM Results

The DAE network is trained to minimize the error between the input and output of the appliance. The training is performed using Stochastic Gradient Descent (SGD) with early stopping to prevent overfitting.

The profile of each household is then created using the preference and need vectors extracted by the NILM technique. Event detection (on/off status) is also used to identify user preferences. The number of on detections of each appliance is calculated at different times of the day, the most frequent ones are identified as user preferences, and the appliance name is added to the preference matrix. After each prediction of the energy consumption of one appliance at a time, the predicted value is compared with the value of the defined threshold to create the need matrix. If the predicted value is more than the threshold, the appliance will be added to the list of high-consumption appliances. A single vector is then

Taghvaei & Safa| Big. Data. Comp. Vis. 1(3) (2021) 161-169

formed from the community of need and preference matrices. A profile is created for each appliance in the secondary database. The weight of each keyword indicating its importance is first calculated using TF-IDF. The weights of the texts are then classified based on the assigned weights. All keywords are sorted in descending order of weight and are labeled by one of eleven appliances.

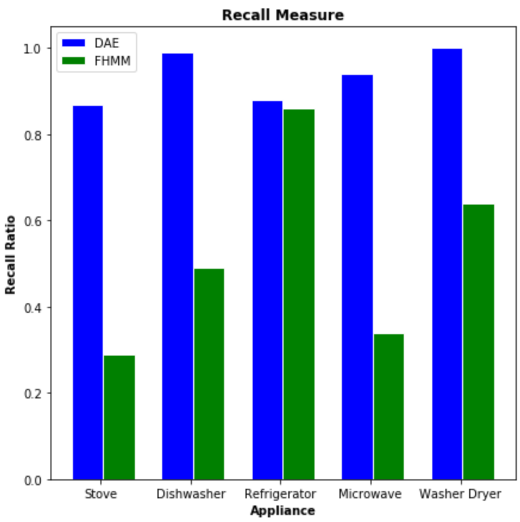
The keywords of the appliance are extracted and matched with all the optimal predicted appliances with the appliance label using cosine similarity (*Eq. (4)*) to match it in a family profile with the optimal one for the recommendation.

$$Score\ (d1.d2) = Cos\ (W_{d1}.W_{d1}) = \frac{W_{d1}.W_{d2}}{\|W_{d1}\|_2.\|W_{d2}\|_2} = \frac{\sum_{i=1}^{k} W_{i.d1}.W_{i.d2}}{\sqrt{\sum_{i=1}^{k} {W_{i.d1}}^2}\ \sqrt{\sum_{i=1}^{k} {W_{i.d1}}^2}} \qquad (4)$$

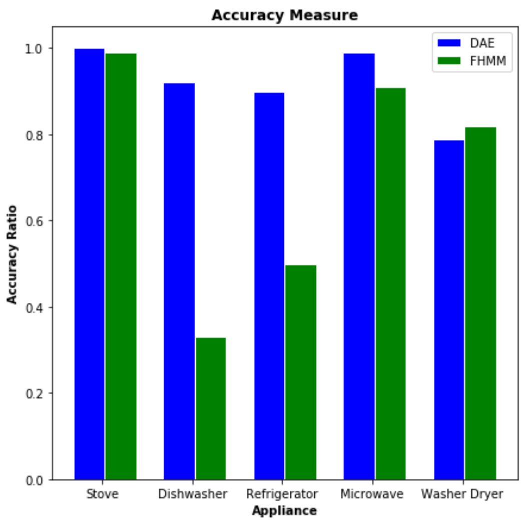
Experimental results

The results of the proposed method (using the Reference Energy Disaggregation Dataset (REDD)[1]) are presented in three sections: 1) Text classification, 2) Energy disaggregation, and 3) Personalized recommendation. The results are obtained with 200 epoch, Adam optimizer, and MSE cost function with Python. In the item profile creation phase, an automated method is used to collect data for various appliances. In this way, about 230 text files for 11 appliances are collected.

The accuracy, precision, recall, F1-score, and relative error values for the NILM model can be seen in *Fig. 3 (A-F)*. As shown, the de-noising auto-encoder performs better than the FHMM. The advantage of deep learning algorithms such as DAE over HMMs is that they can extract high-level features using automated and repetitive learning. The proposed model has a recall of 99%, accuracy, and precision of 80%. Besides, the mean absolute error is 51.27.



(A)

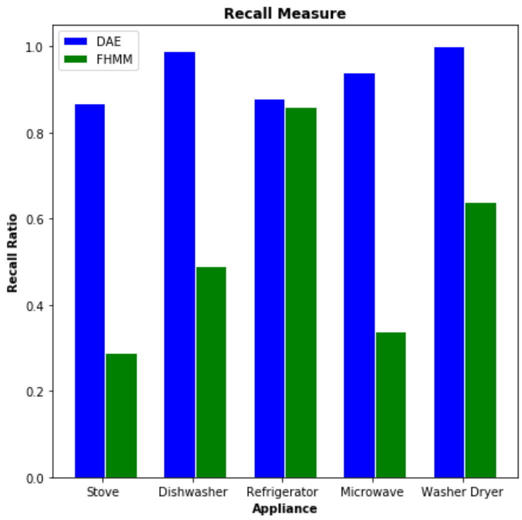


(B)

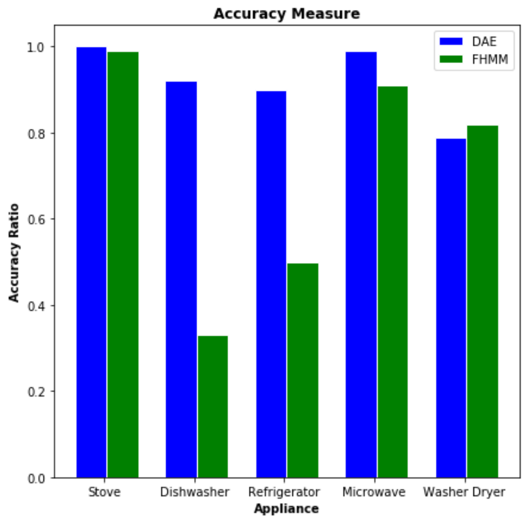
[1] http://redd.csail.mit.edu

Taghvaei & Safa| Big. Data. Comp. Vis. 1(3) (2021) 161-169

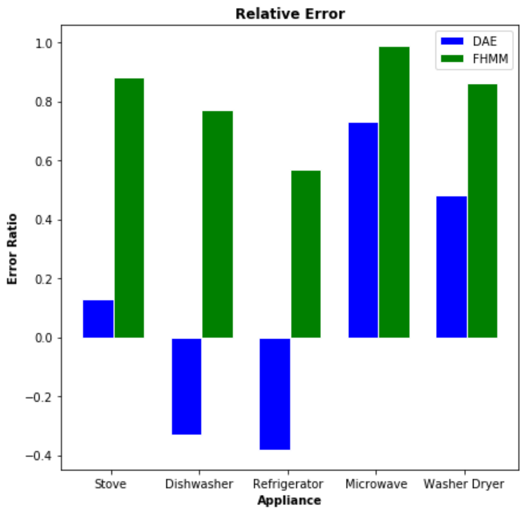
Fig.



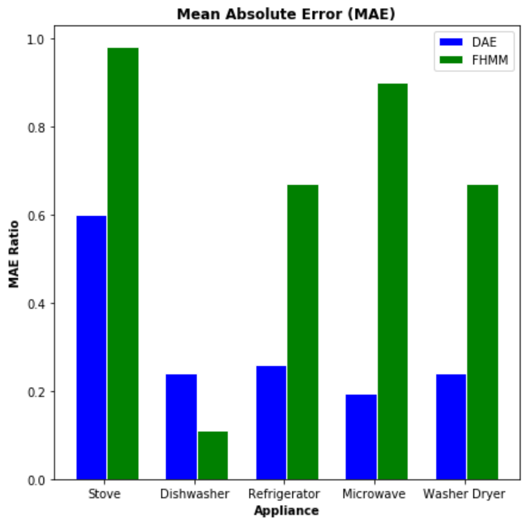
(C)



(D)



(E)



(F)

3. DAE and DHMM Results

The appliances obtained using the proposed NILM and output analysis are listed in *Table 1*. 28 appliances are selected from 68, with 11 selected labels to recommend the optimal ones.

Table 1. List of Appliances

Houses	Appliances
First house	oven_3 — kitchen_outlets_7 — kitchen_outlets_16 — washer_dryer_10 — stove_14 — microwave_11
Second house	kitchen_outlets_3 — kitchen_outlets_8 — Microwave _6 — water_heater_4
Thirth house	washer_dryer_14 — refrigerator_7 — microwave_16 — air_conditioning_14 — air_conditioning_5
Fourth house	air_conditioning_9 — air_conditioning_20 — dishwasher_15— outlets_unknown_6— furnace_4
Fifth house	microwave_3— kitchen_outlets_24— refrigerator_18— furnace_6
Sixth house	air_conditioning_15— kitchen_outlets_13— stove_5— washer_dryer_4

The performance of the TF-IDF algorithm for identifying optimal appliances is shown in *Fig. 4*, where the elements of the main diameter indicate the correct classifications. Out of a total of 230 texts, 185 have one of 11 classes, indicating that the algorithm classifies 80% of the texts as one of the optimal appliances, of which 158 texts are correctly identified, ie 85% accuracy.

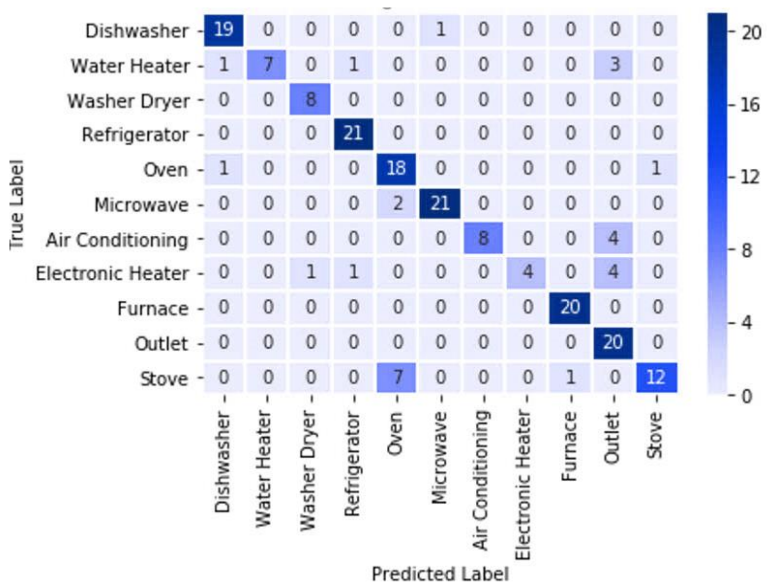


Fig. 4. Confusion matrix for the text classifier

Fig. 5 shows the performance of the recommender system offered for each device in terms of accuracy, recall, and precision (achieved through 230 candidate texts). The results have shown that accuracy for the devices that are organized by the TF-IDF technique is more accurate. Dishwashers, refrigerators, and ovens have the highest rates.

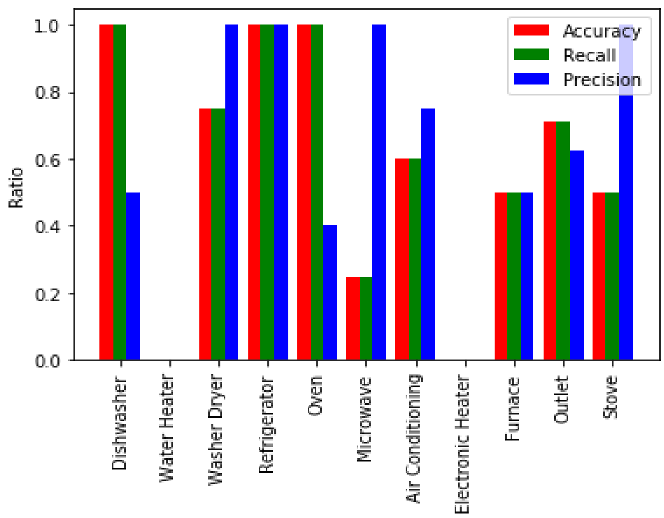


Fig. 5. Performance of the proposed recommender system

The average consumption and recommended values for the 9 appliances are given in *Table 2*. According to the results, the average consumption of refrigerators and dishwashers is higher than the optimal values. On the other hand, the recommended consumption value is more than the value consumed for gas stoves. In general, the proposed method reduces the amount of energy consumption for households by 1234.5 watts.

Table 2. The average consumption and recommended values

Device	Optimal	Consumption	Recommended
Oven	100	110.5	95
Outlets	50	48.6	43.6
Washer Drayer	2500	2130	1860
Stove	1400	1384	1300
Microwave	900	1052	835
Furnace	1000	1235	930
Air Conditioning	2000	2230	2087
Refrigerator	200	225	178
Dishwasher	1500	1750	1600

The average difference between the energy consumption of the recommended appliances compared to the average optimal energy consumption and the average consumption of the appliances in the user profiles is shown in *Fig. 6*. As can be seen in the chart, only air conditioners and dishwashers do not have a better consumption than the average consumption. There are good recommendations for the other seven appliances.

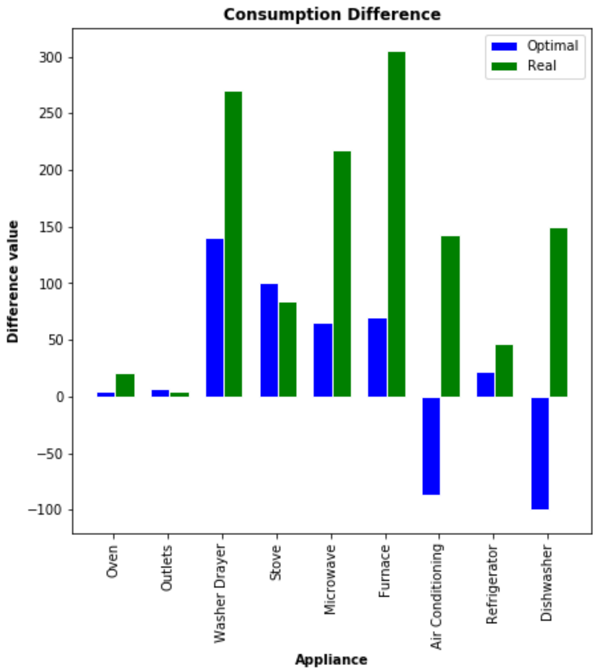


Fig. 6. The average difference between the energy consumption of the recommended appliances.

3 | Conclusions and future directions

This study proposes a personalized TF-IDF and NILM-based recommender system using DAE. The proposed approach was tested using the REDD dataset. Different criteria of accuracy, precision, recall, and F1-score have been used for evaluation. The average NILM results show a performance of 85% and the TF-IDF accuracy for text classification is 80%. In the recommendation phase, the proposed method identifies 28 appliances that have 17 successful recommendations, i.e., 60% accuracy. As a developing field of study, future efforts are recommended to investigate stemming and lexical chains to improve the performance of the TF-IDF technique. Collaborative filtering recommender systems and repetitive learning mechanisms are also important topics that should be addressed in future studies.

Conflicts of Interest

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Cominola, A., Giuliani, M., Piga, D., Castelletti, A., & Rizzoli, A. E. (2017). A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring. *Applied energy*, *185*, 331-344.

[2] Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, *80*(12), 1870-1891.

[3] Piccialli, V., & Sudoso, A. M. (2021). Improving non-intrusive load disaggregation through an attention-based deep neural network. *Energies*, *14*(4), 847.

[4] Yoshimoto, K., Nakano, Y., Amano, Y., & Kermanshahi, B. (2000). Non-intrusive appliances load monitoring system using neural networks. *Growth*, *2*, 2-5.

[5] Kolter, J. Z. N. A., Batra, S., & Ng, A. (2010). Energy disaggregation via discriminative sparse coding. *Advances in neural information processing systems*, *23*, 1153-1161.

Taghvaei & Safa| Big. Data. Comp. Vis. 1(3) (2021) 161-169

[6] Yang, H. T., Chang, H. H., & Lin, C. L. (2007, April). Design a neural network for features selection in non-intrusive monitoring of industrial electrical loads. In *2007 11th International Conference on Computer Supported Cooperative Work in Design* (pp. 1022-1027). IEEE.

[7] Lin, Y. H., & Tsai, M. S. (2010, May). A novel feature extraction method for the development of nonintrusive load monitoring system based on BP-ANN. In *2010 International Symposium on Computer, Communication, Control and Automation (3CA)* (Vol. 2, pp. 215-218). IEEE.

[8] Kelly, J., & Knottenbelt, W. (2015, November). Neural nilm: Deep neural networks applied to energy disaggregation. In *Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments* (pp. 55-64).

[9] Mocanu, E., Nguyen, P. H., Gibescu, M., & Kling, W. L. (2016). Deep learning for estimating building energy consumption. *Sustainable Energy, Grids and Networks*, *6*, 91-99.

[10] Mocanu, D. C., Mocanu, E., Nguyen, P. H., Gibescu, M., & Liotta, A. (2016, October). Big IoT data mining for real-time energy disaggregation in buildings. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 003765-003769). IEEE.

[11] Zhang, C., Zhong, M., Wang, Z., Goddard, N., & Sutton, C. (2018, April). Sequence-to-point learning with neural networks for non-intrusive load monitoring. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).

[12] Deshpande, R., Hire, S., & Mohammed, Z. A. (2022). Smart Energy Management System Using Non-intrusive Load Monitoring. *SN Computer Science*, *3*(2), 1-11.

[13] Kim, H., Marwah, M., Arlitt, M., Lyon, G., & Han, J. (2011, April). Unsupervised disaggregation of low frequency power measurements. In *Proceedings of the 2011 SIAM international conference on data mining* (pp. 747-758). Society for Industrial and Applied Mathematics.

[14] Egarter, D., & Elmenreich, W. (2013, July). Evonilm: Evolutionary appliance detection for miscellaneous household appliances. In *Proceedings of the 15th annual conference companion on Genetic and evolutionary computation* (pp. 1537-1544).

[15] Trung, K. N., Dekneuvel, E., Nicolle, B., Zammit, O., Van, C. N., & Jacquemod, G. (2014, June). Event detection and disaggregation algorithms for nialm system. In *Proceedings of 2nd International Non-Intrusive Load Monitoring (NILM) Workshop*.

[16] Hassan, T. (2012). Bi-level characterization of manual setup residential non-intrusive demand disaggregation using enhanced differential evolution. In *Proc. 1st Int. Workshop Non-Intrusive Load Monitoring*.